



Particulars:

Japanese Patent Laid-Open No. H02-113299

Date of Laid-Open No. April 25, 1990

Japanese Patent Application No. S63-266969

Filing Date: October 22, 1988

Inventors: Hiroya Fujisaki et al

Applicant: Sumitomo Electric Industry Co., Ltd

Agent: Bunji Kamata

1. Title of the Invention:

BASIC FREQUENCY PATTERN GENERATING APPARATUS

2. Claims:

(1) An apparatus using letters, marks and the like as input data, for generating a basic frequency pattern from prosodic features obtained through rules and corresponding tables, being based upon phoneme marks in the inputted data or the inputted data, characterized in that as a phrase command for creating a phrase component among components in the basic frequency pattern, there is provided a means for giving a desired value having a size to be achieved by the phrase component, and for adding or creating a phrase component having a size corresponding to a value which is not yet achieved by the phrase component corresponding to the desired value, as to the phrase command.

9

⑯ 日本国特許庁 (JP)

⑮ 特許出願公開

⑯ 公開特許公報 (A) 平2-113299

⑯ Int.Cl.
G 10 L 5/00

識別記号

府内整理番号
8622-5D

⑯ 公開 平成2年(1990)4月25日

審査請求 未請求 請求項の数 7 (全8頁)

⑭ 発明の名称 基本周波数バタン生成装置

⑮ 特願 昭63-266969

⑮ 出願 昭63(1988)10月22日

⑯ 発明者	藤崎 博也	東京都渋谷区恵比寿3丁目31番12号
⑯ 発明者	山口 幹雄	大阪府大阪市此花区島屋1丁目1番3号 住友電気工業株式会社大阪製作所内
⑯ 発明者	広瀬 啓吉	神奈川県鎌倉市佐助1丁目7番15番
⑯ 発明者	河井 恒	千葉県船橋市前原東5丁目50番8号
⑯ 出願人	藤崎 博也	東京都渋谷区恵比寿3丁目31番12号
⑯ 出願人	住友電気工業株式会社	大阪府大阪市東区北浜5丁目15番地
⑯ 代理人	弁理士 鎌田 文二	

明細書

1. 発明の名称

基本本周波数バタン生成装置

2. 特許請求の範囲

(1) 文字・記号等を入力とし、人力中にある顔様記号によりまたは入力に基づき規則あるいは対応表によって節律的特徴を得て基本本周波数バタンを生成する装置において、

基本本周波数バタンの成分のうちフレーズ成分を作るフレーズ指令として

フレーズ成分の到達すべき大きさの目標値を与え、フレーズ指令に対しては、その目標値に対するフレーズ成分の未到達分に相当する大きさのフレーズ成分を追加または生成する手段を有することを特徴とする基本本周波数バタン生成装置。

(2) 特許請求範囲第1項記載の基本本周波数バタン生成装置において、

フレーズ成分が対数軸上の基本本周波数バタンにおいて、臨界制動二次線形系のインパルス応答の和で近似することを特徴とする基本本周波数バタン生

成装置。

(3) 特許請求範囲第2項記載の基本本周波数バタン生成装置において、

臨界制動二次線形系のインパルス応答の式を

$$G_p(t) = \begin{cases} -\alpha^2 t \exp(-\alpha t) & (t \geq 0) \\ 0 & (t < 0) \end{cases}$$

とおくとき、
1番目のフレーズ指令の目標値と、1番目のフレーズ指令の時点から t/α 後の時点における、 $t - 1$ 番目までのフレーズ指令によるフレーズ成分と、

の差をフレーズ成分の未達分とすることを
を特徴とする基本本周波数バタン生成装置。

(4) 特許請求範囲第1項記載の基本本周波数バタン生成装置において、

フレーズ成分の未達分が負の場合、すなわち、フレーズ成分の追加なくフレーズ指令の目標値に達する場合は、フレーズ指令の大きさを0とする、すなわち、フレーズ指令を生成しないことを
を特徴とする基本本周波数バタン生成装置。

(5) 特許請求範囲第1項記載の基本本周波数バタン

特開平2-113299 (2)

生成装置において、

直前のフレーズ指令の大きさと時間間隔を基にしてテーブルによりフレーズ成分の大きさを求め、その値によって

追加のフレーズ指令の目標値に対するフレーズ成分未進分を求める。

フレーズ指令の大きさを定めることを特徴とする基本周波数バタン生成装置。

(6) 特許請求範囲第1項記載の基本周波数バタン生成装置において、

直前のフレーズ指令の目標値と時間間隔と追加のフレーズ指令の目標値とを基にしてテーブルによって追加のフレーズ指令の大きさを定めることを特徴とする基本周波数バタン生成装置。

(7) 特許請求範囲第1項記載の基本周波数バタン生成装置において、

フレーズ指令に対しては、その目標値とそのフレーズ指令の時点におけるフレーズ成分との差によって追加または生成するフレーズ指令の大きさを

求めることにより、

フレーズ成分の追加または生成を行うことを特徴とする基本周波数バタン生成装置。

3. 発明の詳細な説明

(産業上の利用分野)

本発明は、音声合成装置、特に規則により基本周波数バタンを生成する音声規則合成装置やテキスト合成装置に用いられる基本周波数バタン生成装置に関するものである。

(従来の技術)

文字・記号等を入力とした従来の音声合成装置(たとえば、昭和61年度電子通信学会発行全国大会講演文集S26-5、昭和61年3月)の処理プロック図を第2図に示す。また上記音声合成装置への入力例を第1図に示す。

入力は、アクセント記号、フレーズ記号、休止記号及び音節記号である。

アクセント記号はA1(0.40)、A2(0.26)及びAOであり、A1とA2によってアクセントの立上がりの音節境界と、アクセントの大きさの種類を

示し、AOでアクセントの立下りの音節境界を示す。

アクセントの大きさは、A1、A2の記号で示される2種類がある。()内は実際に割当てるアクセントの大きさを示す。

フレーズ記号は、P1(0.43)、P2(0.26)、P3(0.12)及びPOであり、P1、P2、P3によってフレーズの開始時点と大きさの種類を示し、POはそれ以前にあるP1、P2、P3によって生じたフレーズ成分を0に減らすことを示している。()内は割当てるフレーズの大きさを示す。

休止記号は「。」(0.7秒)、「、」(0.3秒)「・」(0.08秒)であり、その音節境界で休止がおかれること、すなわち間があくことを示している。()内は休止の長さを示す。

音節記号は「ス」「ズ」「メ」「ワ」……等の片仮名で表された記号であり、音の種類を示している。

入力中、○のついたもの、たとえば「シ○」は無声化された「シ」を示す。

第一回

```

P1スAIズメKA0ライPO、
PIAIダAOレデモ
P3シ○A2カティルトリワAOアA2ルマAOイPO。
PIムAIカシバAOナシヤ
P3コA2トワザニAOモ・
P2トA2ージ・スルAOノダカラPO、
P1スA2ズノワAOPO、
PIAIカAOナリA2イAOゼンカラ・
P2ニAIホンジAOシニ
P3シ○AIカシマAOレテキ○A2タAOノダローゴーPO。

```

入力によって合成すべき音声が指示されるが、その合成処理は次のようにして行なう。

(1) 音素的バラメータの生成

音節記号により指定される音節を蓄積バタンから選び出し、その蓄積バタン中に記載されているその音節固有の時間長と、休止記号によって指定される休止時間をもとにしてその音節の時点を決定する。

次に、蓄積バタンに記載されているその音節固

特開平2-113299(3)

有の音素的パラメータ、例えばホルマント周波数と帯域幅の時間変化バタンを読み出し、先に定めた各音節の時点が満たされるように、音素的パラメータを伸長・圧縮させながら、互いにつなぎあわせる。たとえば、ある音節 $C_1 V_1$ の時点を $t = 0$ とし、次の音節 $C_2 V_2$ が $t = 140\text{msec}$ であり、 $C_1 V_1$ の音素的パラメータが $t = 100\text{msec}$ の分までしか記述されておらず、また $C_1 V_1$ の音素的パラメータが $t = 140\text{msec}$ からの分からしか蓄積バタンに記述されていないとしたら、 $t = 100\text{msec}$ から $t = 140\text{msec}$ までの間は、 V_1 の部分を引き伸ばすことにより揃う。

以上の処理によって、合成しようとする文音声の音素的パラメータが得られ、音声合成器（たとえばホルマント合成器）に送られて、音声信号生成に用いられる。

(2) 音源強度バタンの生成

音源強度は、合成する音節の種類ごとに値を定め、また休止の前・後では、減少・増加させる必要がある。音節の種類に固有な値は、やはり蓄積

バタンに記載されており、音素的パラメータと同様の伸縮処理を行なってつなぎ合わせることで、目的とする文の基礎的な音源強度バタンが得られる。さらに、休止、特に文と文の区切りを示す「。」の休止の前後で、音源強度規則にしたがって、一定量を減少・増加させることで、最終的な音源強度バタンが得られ、音声合成器に送られて音声信号生成に用いられる。

(3) 基本周波数（声の高さ、記号 F_0 です）バタン生成

入力中には、フレーズ及びアクセントの時点が、どの音節境界にあるかが示されており、しかも前述のように音節の時点が決定しているので、音節の時点を基準としてフレーズおよびアクセントの時点を決められる。また入力中のフレーズ記号、アクセント記号の種類によって実際に用いる値が決まっている（たとえば、 $A1$ は 0.40 ）ので、これによりフレーズ指令とアクセント指令の大きさと時点を決めることができる。フレーズおよびアクセントの時点と大きさをもとにして F_0 バ

タンの生成モデルの式により、 F_0 バタンの生成を行なう。

F_0 バタン生成モデルを第3図に示す。 F_0 の時間変化バタンを $F_0(t)$ で表し、次の式によって計算を行う。

$$\begin{aligned} -\ln F_0(t) &= \ln F_{min} + \sum_{i=1}^I AP_i G_p(t - T_{pi}) \\ &+ \sum_{j=1}^J AA_j (G_a(t - T_{aj}) - G_a(t - T_{ej})) \dots \textcircled{1} \\ G_p(t) &= \begin{cases} \alpha^2 t \exp(-\alpha t) & (t \geq 0) \\ 0 & (t < 0) \end{cases} \\ G_a(t) &= \begin{cases} \min(1 - (1 + \beta t) \exp(-\beta t), \theta) & (t \geq 0) \\ 0 & (t < 0) \end{cases} \end{aligned}$$

であり、それぞれ臨界制動二次線形系のインバルス応答と、ステップ応答になっている。

α 、 β は応答の速さを決める定数であり、 $\alpha = 3.0$ 、 $\beta = 20.0$ 程度の値を用いる。

$1 - (1 + \beta t) \exp(-\beta t)$ は、 t が増加するに伴って目標値 1.0 に漸近するが、有限の時間内で $G_a(t)$ を目標値に収束させるため、 $\theta = 0.9$ として処理を行っている。 $\theta \leq 1$ の条件の場合 $G_a(t)$ の目標値は θ である。

「」は、その文章内に出てくるフレーズの数を示し、 Api は i 番目で示されるフレーズ指令の大きさを示す。たとえば、 $P1 (0.43)$ の記号で示されるフレーズが来るなら $Api = 0.43$ となる。 T_{pi} はそのフレーズ指令の時点を示す。

「」は、その文章内に出てくるアクセントの数を示し、 AAj は j 番目で示されるアクセント指令の大きさを示す。たとえば、1番目のアクセントとして $A1 (0.40)$ の記号で示されるアクセントが来るならば、 $AA1 = 0.40$ となる。 T_{aj} 、 T_{ej} は、 j 番目のアクセント指令の開始時点と終了時点を示す。

$InPain$ は、定数項であり、声帯の振動可能最低周波数に対応している。たとえば、男性音声を合成するときは、 $F_{min} = 75\text{Hz}$ 程度に、女性音声を合成するときは $F_{min} = 115\text{Hz}$ 程度に設定する。

$F_0(t)$ を計算するときは、前述の処理によって決まったフレーズ指令の大きさと時点 Api 、 T_{pi} ($1 \leq i \leq I$)、アクセント指令の大きさと時点 AAj 、 T_{aj} 、 T_{ej} ($1 \leq j \leq J$) を前掲の①式に

特開平2-113299 (4)

当てはめて右辺を計算し、その結果に対して対数の逆関数、すなわち指数関数をとることにより、 $P_e(L)$ を計算する。

以上の処理によって得られた $P_e(L)$ 、すなわち基本周波数バタンは、音声合成器に送られ、音声信号の生成に用いられる。

上述の処理に用いられるハードウェアは、音声合成器（たとえばホルマント合成器）は、信号処理プロセッサにより実現されており、入力記号から音声合成器への入力を作成するまでの処理は、マイクロプロセッサによって処理される。蓄積バタンは、マイクロプロセッサのアクセスする ROM に記憶される。

また、 $P_e(L)$ の計算式の計算などはマイクロプロセッサのプログラムで実現されている。

以上の説明では、音節記号、休止記号、アクセント記号、フレーズ記号を入力して、音声を合成する場合の処理を示したが、漢字仮名混じり文章を入力する音声合成装置も知られている。この場合、漢字仮名混じり文を前述した音節記号、休止

記号、アクセント記号、フレーズ記号に変換する処理が必要である。この処理は、入力文章を単語単位に分かち、単語辞書を検索することにより読みを決定し、同じく単語辞書に書かれたアクセント型により、アクセントの上り下りの音節境界を決定し、アクセントの大きさを割り当てる処理によって行なわれる。

次に、従来技術による休止記号・フレーズ記号を合成しようとする文への与え方を説明する（前出論文参照）。

この与え方は、文の統語構造（構文）に基づいており、次の通りである。なお、統語構造から休止記号・フレーズ記号を導く導出規則は、これ以外にも種々の変形・改良版がある。

- (1) 文の句点に“。”と P0、文頭に P1 を置く。
- (2) 文の誤点に“、”と P0 と P1 を置く。
- (3) 読点（なければ句点）と比較して統語的に大きな語境界、または 1 段小さな語境界には“・”と P2 を置く。
- (4) 読点と比較して 2 段小さな語境界には P3 を置

く。

(5) ただし修飾関係にある語の境界では前記(3)(4)にかかわらず休止・フレーズ記号をおかない。

(6) 以上のようにして設定した休止・フレーズ記号の間隔がある程度以上離れていた場合（通常の発話速度で 13 モーラ程度）、大きい語境界の順に P3 を追加する。

(7) 統語上の境界を特に示す必要がある場合には小さな語境界でも P3 を置く。

(8) すべての P2 について直前の P1、P2 との間隔をしらべ、間隔が小さければ（通常の発話速度で 4 モーラ程度）P3 を変更する。

〔発明が解決しようとする課題〕

①式によりフレーズ成分を計算する場合、フレーズ指令が短い時間間隔で続くと、前のフレーズ指令によるフレーズ成分が大きく残っている間に次のフレーズ指令によりフレーズ成分を追加することになり、フレーズ成分全体としての大きさはかなり大きくなる。ところが自然音声の発話では、声の高さは生理的制約によりむやみに高くなるこ

とはない。このため、フレーズ指令が短い時間間隔で続く場合は、前記⑧の処理により、フレーズ指令の大きさを小さくする必要があった。

すなわち、フレーズ指令自体は本質的には文の統語構造を反映して定まると考えられるが、実際に与えるフレーズ指令の大きさはそれまでのフレーズ指令により修正する必要がある。そのため、文の統語構造とフレーズ指令の大きさは対応が明確ではなくなり、フレーズ指令の与え方の規則は見通しが悪くならざるを得なかった。

本発明は、この生理的制約に相当するフレーズ指令大きさ修正処理をフレーズ成分生成処理に内在させ、もって見通しのよい統語生成規則を可能とすることを目的としている。

〔課題を解決するための手段〕

本発明は、フレーズ指令としてフレーズ成分の目標値を与え、フレーズ指令の大きさはその目標値に達するために要するフレーズ成分の大きさから求めるこことを特徴とする。

〔作用〕

特開平2-113299 (5)

第1図を用いてフレーズ指令の大きさの求め方を説明する。

i番目のフレーズ指令として、指令の時点 t_0 と目標値 T_{p_i} が与えられる。一方フレーズ制御規格2により、i-1番目迄のフレーズ指令によるフレーズ成分が定められる。

本発明はi-1番目迄のフレーズ指令によるフレーズ成分と、i番目のフレーズ成分によって達成されるべきフレーズ成分の目標値 T_{p_i} との差を目標未達分計算部1により計算し、その差に相当するフレーズ成分を生成するのに要するフレーズ指令の大きさ A_{p_i} を求める。

〔第1実施例〕

フレーズ成分の正の目標値として2通りある場合の実施例を説明する。

①式における $G_p(l)$ は、 $l = 1/\alpha$ において最大値 $G_p(1/\alpha) = \alpha/e$ (e は自然対数の底) を取る(第2表)。そこで、従来技術でのPIに対するフレーズ指令の大きさとして0.43を割り当てていたのを、 e/α 倍して、第1のフレーズ成

分の目標値は $0.43 \times e/\alpha$ とする。 $\alpha = 3.0$ 、 $e = 2.71828$ とすると、フレーズ成分の目標値は0.47となる。第2のフレーズ成分の目標値として、第1の目標値の8割すなわち、 $0.47 \times 0.8 = 0.376$ とする。

次に、フレーズ成分の目標未達分の計算実施例を第4図(4)を用いて説明する。i-1番目迄のフレーズ指令によるフレーズ成分に対し、時刻 $l = 0$ でi番目のフレーズ指令が発生してフレーズ成分が追加される場合、フレーズ指令が極大値を取る時間 t_{max} は $1/\alpha$ と異なる。しかし、 t_{max} は簡単な計算では求めないこと、感覚上はフレーズ指令の大きさは必ずしも厳密に制御する必要はないこと、を考慮して、 $l = 1/\alpha$ におけるフレーズ成分未達分をi番目のフレーズ指令の大きさの計算に用いることとする。すなわち、第4図(4)では、 $T_{p_i} - c$ がフレーズ成分未達分である。

最後に、フレーズ指令の大きさの計算実施例を説明する。フレーズ指令の大きさは、第4図(4)より、フレーズ成分未達分の e/α 倍とすればよい。

そこで、フレーズ成分未達分が $T_{p_i} - c$ とすると、与えるべきフレーズ指令の大きさは、 $(T_{p_i} - c) \times e/\alpha$ となる。ただし、与えるべきフレーズ指令の大きさが負の値になった場合は、フレーズ指令の大きさは0とする(i番目のフレーズ指令を生成しない)。

なお、負のフレーズ記号 P_0 は、フレーズ成分の下がりを実現するためなので、従来技術と同様に-0.5を固定的に割り当ててもよく、また、フレーズ成分の目標値として0を割り当てて、正のフレーズ指令と同様にしてフレーズ指令の大きさを求めてよい。

〔第2実施例〕

感覚上は、フレーズ指令の大きさは厳密に制御する必要がない点に着目し、第1実施例よりも計算が簡単な実施例を次に説明する。

i-1番目迄のフレーズ指令によるフレーズ成分を求めるときは、 $G_p(l) = \alpha^2 t \exp(-\alpha t)$ の計算を行う必要があるが、関数は計算時間がかかる。そこで、これを表に記載しておいて検索す

れば、計算時間を短縮できる。そして、フレーズ指令の大きさが厳密に欠けても感覚上差し隠りがないので、この表としては比較的粗い時間間隔(たとえば0.1秒)で記載することで記憶容量を減らすことができる(第2表)。

さらに、 $G_p(l)$ の関数は l の増大とともにむずかしくに0に漸近するので、i-1番目迄のフレーズ指令によるフレーズ成分はi-1番目のフレーズ指令によるフレーズ成分が主であり、i-2番目迄のフレーズ指令による寄与分は少ない(第5図)。

以上の観点から、i-1番目のフレーズ指令の大きさを用いて、i番目のフレーズ指令の大きさを求める実施例を次に示す。

まず、第1実施例と同様に、i番目のフレーズ指令の時点から $1/\alpha$ 後のフレーズ成分の大きさを求める。この時点のフレーズ指令の大きさは、i-1番目のフレーズ指令の時点からの経過時間で第2表を検索して $G_p(l)$ の値を求め、i-1番目のフレーズ指令の大きさを掛け合わせること

で求め、この値を以降は c と書く。

第2表

t の範囲	$G_p(t)$ の値
0 $\leq t < 0.1$	$G_p(0.05) = 0.39$
0.1 $\leq t < 0.2$	$G_p(0.15) = 0.86$
0.2 $\leq t < 0.3$	$G_p(0.25) = 1.06$
0.3 $\leq t < 0.4$	$G_p(0.35) = 1.10$
0.4 $\leq t < 0.5$	$G_p(0.45) = 1.05$
0.5 $\leq t < 0.6$	$G_p(0.55) = 0.95$
0.6 $\leq t < 0.7$	$G_p(0.65) = 0.83$
0.7 $\leq t < 0.8$	$G_p(0.75) = 0.71$
0.8 $\leq t < 0.9$	$G_p(0.85) = 0.60$
0.9 $\leq t < 1.0$	$G_p(0.95) = 0.49$
1.0 $\leq t < 1.1$	$G_p(1.05) = 0.40$
1.1 $\leq t < 1.2$	$G_p(1.15) = 0.33$
1.2 $\leq t < 1.3$	$G_p(1.25) = 0.26$
1.3 $\leq t < 1.4$	$G_p(1.35) = 0.21$
1.4 $\leq t < 1.5$	$G_p(1.45) = 0.17$
1.5 $\leq t < 1.6$	$G_p(1.55) = 0.13$
1.6 $\leq t < 1.7$	$G_p(1.65) = 0.11$
1.7 $\leq t < 1.8$	$G_p(1.75) = 0.08$
1.8 $\leq t < 1.9$	$G_p(1.85) = 0.06$
1.9 $\leq t < 2.0$	$G_p(1.95) = 0.05$
2.0 $\leq t$	0

次に、フレーズ成分未達分を求める $T_{p_i} - c$

2通りある場合の A_{p_i} の値を種々の i に対して表に示すと第3表の通りとなる。実際にフレーズ指令を与える場合、 A_{p_i} が負の場合 $A_{p_i} = 0$ とする、すなわち、フレーズ指令は生起しない。直前のフレーズ指令によるフレーズ成分は、徐々に減衰するので、 $i > 20$ の場合は $i = 20$ と同じ扱いにすればよい。

第3表

i の値	$T_{p_i} = 0.47$		$T_{p_i} = 0.376$	
	$T_{p_i} = 0.47$	$T_{p_i} = 0.376$	$T_{p_i} = 0.47$	$T_{p_i} = 0.376$
1	0.023	0.104	-0.062	0.019
2	0.074	0.144	-0.011	0.059
3	0.132	0.191	0.047	0.106
4	0.188	0.236	0.103	0.151
5	0.233	0.276	0.153	0.190
6	0.280	0.303	0.195	0.224
7	0.314	0.336	0.229	0.251
8	0.341	0.358	0.256	0.273
9	0.362	0.375	0.277	0.290
10	0.378	0.388	0.293	0.303
11	0.391	0.398	0.305	0.312
12	0.400	0.405	0.315	0.320
13	0.407	0.411	0.322	0.325
14	0.412	0.415	0.327	0.330
15	0.416	0.418	0.331	0.333
16	0.418	0.420	0.333	0.335
17	0.421	0.422	0.335	0.336
18	0.422	0.423	0.337	0.338
19	0.423	0.424	0.338	0.338
20~	0.424	0.424	0.339	0.339

特開平2-113299 (6)

となる。最後に、1番目のフレーズ指令の大きさは第1実施例の場合にして $(T_{p_1} - c) \times e / \alpha$ とする。

(第3実施例)

第2実施例では、直前のフレーズ指令の大きさをもとにフレーズ成分未達分を求めており、直前のフレーズ指令の目標値をもとにフレーズ成分未達分を近似することもできる。

さらに、直前のフレーズ指令からの時間間隔として拍数によってカウントすることもできる。

これらの点を考慮した実施例を次に示す。

発話速度を m 拍/秒とすると、1拍はなれた時間間隔は $1/m$ である。直前のフレーズ指令の目標値 T_{p_i} に対して、 $T_{p_i} \times e / \alpha$ の大きさのフレーズ指令が生じているとする。1拍はなれた次のフレーズ指令の目標値 $T_{p_{i+1}}$ に対する未達分 d は $d = T_{p_{i+1}} - T_{p_i} \times e / \alpha \times G_p(1/m + 1/\alpha)$ となる。よって次のフレーズ指令の大きさは、 $A_{p_{i+1}} = d \times e / \alpha$ となる。

フレーズ指令の大きさとして 0.47 と 0.376 の

(第4実施例)

第3実施例は、フレーズ指令の大きさとして $(T_{p_1} - c) \times e / \alpha$ の値をそのまま用いているが、フレーズ指令の大きさを何段階かに量子化を行うことも可能である。すなわち、フレーズ指令の大きさとして、0.43、0.26、0.12 の3種類（それぞれ、P1、P2、P3と記号を付ける）用意しておき、 A_{p_i} に最も近い値を採用することもできる。

この観点から第3表を書き直したものを作成する。この表によれば、極めて簡単にフレーズ指令の大きさを決めることができる。

第4表

i の値	$T_{p_i} = 0.47$		$T_{p_i} = 0.376$	
	$T_{p_i} = 0.47$	$T_{p_i} = 0.376$	$T_{p_i} = 0.47$	$T_{p_i} = 0.376$
1	0	P_1	0	P_1
2	P_1	P_2	0	P_2
3	P_2	P_3	0	P_3
4	P_3	P_1	P_3	P_1
5	P_3	P_2	P_3	P_2
6	P_2	P_1	P_2	P_1
7	P_2	P_1	P_2	P_1
8	P_1	P_1	P_1	P_1
9~	P_1	P_1	P_1	P_1

特開平2-113299 (7)

(第5実施例)

第1実施例では、1番目のフレーズ指令の時点から ϵ/α 時間後の、1-1番目迄のフレーズ指令によるフレーズ成分の大きさ c を必要としていた。処理上の簡便さの点では、将来の時点の値を予想した値を使うよりも現在までに判明している値で処理が行える方が簡便である。そこで、1番目のフレーズ指令の時点で得られるフレーズ成分の大きさ c' を、 c の代わりに用いる実施例を次に説明する。

まず、1番目のフレーズ指令の時点 T_{p1} において、フレーズ成分の値 c' を求める。音声波形生成のために基本周波数バタンを求める処理を行っているので(第2図参照)、時点 T_{p1} において、フレーズ成分の値 c' を取り出せばよい(第6図)。

次に、フレーズ成分未達分を求める $T_{p1} - c'$ となる。最後に、1番目のフレーズ指令の大きさは第1実施例の場合と同様にして $(T_{p1} - c') \times \epsilon/\alpha$ とする。

(発明の効果)

本発明によれば、フレーズ成分が大きくなり過ぎない様にするための、フレーズ指令の大きさ設定処理が、フレーズ成分生成処理に内在化することができる。このことにより、従来は必要であったフレーズ指令の大きさ修正処理が不要になり、フレーズ指令生成規則が簡単になり見通しが良くなる。

4. 図面の簡単な説明

第1図は、本方式のフレーズ指令の大きさ算出処理概念図

第2図は、音声合成装置の処理ブロック図

第3図は、F.バタン(基本周波数バタン)生成モデル

第4図の(同)及び(例)は、追加のフレーズ成分説明図

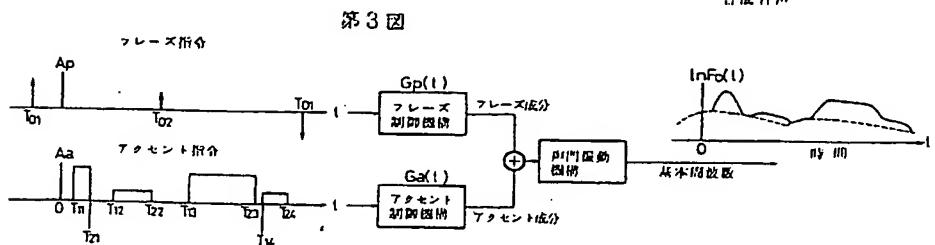
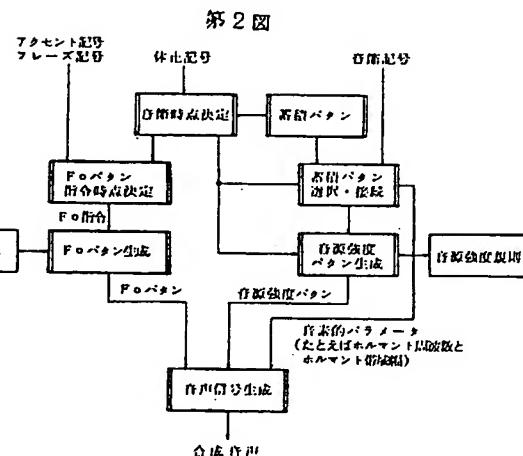
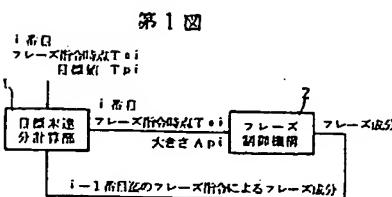
第5図は、フレーズ成分の大きさ算出説明図

第6図は、フレーズ成分未達分簡略計算法の説明図

である。

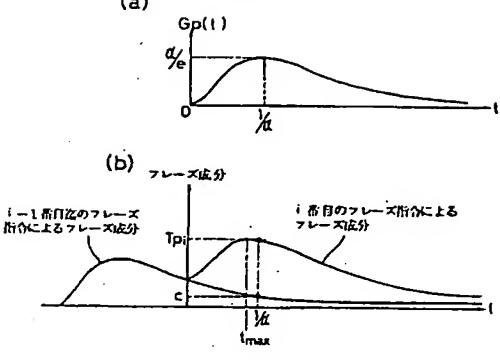
1……目標未達分計算部、

2……フレーズ制御機構。

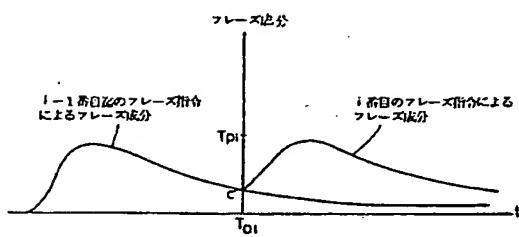


特開平2-113299 (8)

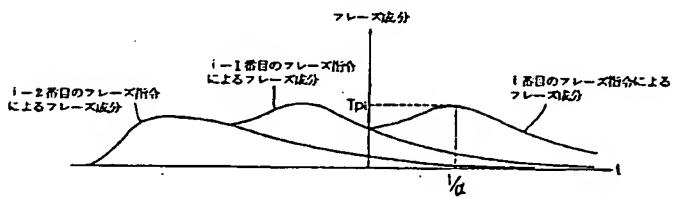
第4図



第6図



第5図



【公報種別】特許法第17条の2の規定による補正の掲載

【部門区分】第6部門第2区分

【発行日】平成8年(1996)10月18日

【公開番号】特開平2-113299

【公開日】平成2年(1990)4月25日

【年通号数】公開特許公報2-1133

【出願番号】特願昭63-266969

【国際特許分類第6版】

G10L 5/00

【F1】

G10L 5/00 8946-5H

手 続 拡 正 書 (自効)

平成7年7月25日

特許庁長官原

1 事件の表示

昭和63年特許願第266969号

2 発明の名称

基本周波数パタン生成装置

3 補正をする者

事件との関係 特許出願人

住所 東京都渋谷区恵比寿8丁目31番12号

氏名(名称) 藤崎博也

(平成元年2月13日付
〔行政取扱事項による住所変更届出〕)

住所 大阪市中央区北浜四丁目5番83号

氏名(名称) 住友電気工業株式会社

4 代理人

住所 〒542 大阪市中央区日本橋1丁目18番12号

氏名 (7420)井理士 錦田文二

Tel. 06(631)0021 Fax. 06(641)0024

5 拡正の対象

明細書の「発明の詳細な説明」の範

6 拡正の内容

明細書の第9頁下から第8行目にある「テスラ応答」を「ステップ応答」に補正します。

特許